International Journal of Advanced Engineering, Management and Science (IJAEMS)
Peer-Reviewed Journal
ISSN: 2454-1311 | Vol-10, Issue-3; Mar-Apr, 2024
Journal Home Page: https://ijaems.com/
DOI: https://dx.doi.org/10.22161/ijaems.103.5

Guardians of E-Commerce: Harnessing NLP and Machine Learning Approaches for Analyzing Product Sentiments in Online Business in Nigeria

Mbeledogu Njideka Nkemdilim, Ogbu Ikechukwu Michael

[1]Department of Computer Science, Nnamdi Azikiwe University, Awka, Nigeria

Received: 02 Feb 2024; Received in revised form: 18 Mar 2024; Accepted: 29 Mar 2024; Available online: 06 Apr 2024

***Abstract*— *In today's e-commerce in Nigeria, customers access online stores to browse through and place orders for products or services via the internet on their devices while some are skeptical due to the experiences from what I ordered versus what I got syndrome. Though this method of business has flourished to an extent, it greatly faces a crucial challenge in unravelling consumer's sentiments particularly in the realm of product reviews. This deficiency inhibits most e-commerce platforms in Nigeria from gaining effective sensitivity into users' preferences, thus, limiting their ability to boost their product recommendations and, understand and improve customers' experiences. This research aims to bridge this gap by developing a sentiment analyzer of product in the e-commerce domain using Natural language processing and machine learning approach. The model will analyze the customers' reviews based on positive or negative. The experimental data was collected from kaggle.com. Stemming and lemmatization were approaches used for cleaning the collected data. Features were extracted and transformed using CountVectorizer. Gaussian Naïve Bayes classifier was used as the machine learning technique. The model's performance was evaluated and it returned 90% of accuracy, hence, an efficient and reliable model for product review sentiment analysis is developed.*

***Keywords*— E-commerce, Natural language processing, Sentiment analysis, Machine learning**

I. INTRODUCTION

Commerce as the engine that drives the economy of societies by promoting business inventions has brought about advanced growth in the nation. It has developed from the trade-off systems to a more global network of businesses from the inception of the internet over the years. This act of encompassing commerce by the internet gave birth to electronic commerce. Electronic commerce (e-commerce) is the buying and selling of goods and services, or the transmitting of funds or data over an electronic network, primarily the internet (Lutkevich *et al.*, 2023) via devices. This transaction occurs either as business-to-business (B2B), business-to-consumer (B2C), consumer-to-consumer (C2C), consumer-to-business (C2B), business-to-administration (B2A), consumer-to-administration (C2A) or mobile e-commerce (m-commerce).

No matter the type of e-commerce a consumer delves into, the benefits are easy and fast accessibility, wide visibility of goods and services, and 24/7 hours market service availability. Despite these advantages, it is still greatly challenged in the areas of cybersecurity, competition, order fulfillment, customer's experience, quality website traffic and visitor conversion, visibility, return and refund policies, finding the right market, making and increasing sales, borderless e-commerce and augmented reality (Post, 2023). Amongst these challenges, the most troubling issue is customer's experience. Presently, customers that would like to order products online are skeptical in doing such because of the experiences many encountered with ordering products online. One of such is receiving an item not in the form of what was initially shown to the buyer and not having the opportunity to lay compliant in that regards. This syndrome was tagged with a slogan termed "what I ordered versus what I got".

In order to find the product-market fit, the reviews of the customers are very much needed. It is the first step of any

business, and e-commerce is not different. It is an important metric that helps a business manager address customer's issue, prevent churn and build a base of loyal customers (Needle, 2023). One will learn what the customers like about their offers and areas where they can improve, thereby, providing a method of facilitating development. This development involves personal growth, motivation, creativity, morale, job satisfaction, clarity, professional relationships, meaningful discussion and goal alignment. It is important to note that in absence of a product-market fit, some of the challenges such as competition and poor sales of products are as a result of the shortcomings of customer review.

In the light of this grave challenge, there is a need to create the so much needed avenue that can handle the review of customers in order to boost productivity in business.

II. LITERATURE REVIEW

Artificial Intelligence (AI) which is sometimes called machine intelligence is the intelligence demonstrated by machines in contrast to the natural intelligence displayed by humans and other animals (Ziyad, 2019). It handles the important question of what knowledge is needed in any mode of thinking and how should that knowledge be applied. The basic components of AI are learning, reasoning and decision making, problem solving and perception. There are six main branches of AI. They are machine learning, deep learning, expert systems, Natural Language Processing (NLP), Robotics and Fuzzy logic (Tyagi, 2020). Amongst these branches, the one that addresses interpretation, comprehension and manipulation from human language in order to resolve the issue of customers' review is the NLP.

Natural Language Processing

Natural Language Processing is a subset of Artificial Intelligence that facilitates seamless communication between humans and computers by granting machines the capacity to interpret, comprehend, generate, manipulate and extract significance from human languages (Oracle, 2022). It has become a crucial area of research and development that merges interdisciplinary fields in computer science, linguistics, and cognitive psychology. Due to the available large volume of text and voice data in various forms, such as posts from online or social media and news articles, NLP covers different tasks such as text analysis, speech recognition, language translation, text summarization, and sentiment analysis.

Sentiment Analysis

Sentiment analysis is literally known as opinion mining. It is the process of determining the emotional tone behind a piece of text, whether it is positive, negative, or neutral (Biyani 2023). Naturally, companies receive feedbacks from countless sources, including customer messages, call center and social media posts. These sources are the different views of people about the companies' products. Based on the analysis of such sources, the companies gain valuable insights in their businesses as they identify areas for improvement in terms of products and rendered services, thereby, creating and promoting the overall satisfaction for their customers.

In situations where very large volume of textual data are collected, the act of manually analyzing these textual data is not feasible. In order to effectively analyze such, statistical, natural language processing and machine learning techniques are applied to determine the emotional meaning of communications (Korolov, 2021). Statistical methods are used to estimate and evaluate the occurrences of positive, negative, or neutral sentiments in a given dataset, machine learning algorithms (supervised) train the labeled datasets to recognize the available patterns and make predictions while NLP techniques enables the system to have an extensive acumen of the emotional context within the text of the human language by identifying words, phrases and linguistic constructs that are sentiment based. Sentiment analysis automates the process of analyzing large volumes of text data and extracting feelings, emotions and attitudes about a product or services generated by depending on NLP, thereby, saving time and effort when compared to manual analysis.

Review of Related Works

Numerous studies have explored the use of sentiment analysis for product in e-commerce. Some of them are:

Alharbi *et al*. (2021) conducted a research aimed at accurately forecasting customer feedback based on smartphone reviews gathered from Amazon.com. Deep Learning (DL) methods were employed in this study, focusing on the analysis and categorization of reviews into three classes: positive, neutral, and negative. The authors explored various DL approaches including Recurrent Neural Networks (RNN) and its four variants—Gated Recurrent Unit (GRU), Long Short Term Memory (LSTM), Update Recurrent Neural Netwok (UGRNN), and Group Long Recurrent Neural Network (GLRNN). These algorithms were complemented by word embedding techniques, utilizing Glove, word2vec, and FastText as feature extraction approaches.

Norrega *et al*. (2023) delved into a comprehensive investigation into Amazon reviews, employing advanced deep learning techniques. Alongside the primary aim of the research, a randomly relatable objective was to explore the impact of sentiment analysis on customer satisfaction in e-

commerce. The models utilized in this study encompassed Bidirectional Encoder Representations from Transformers. (BERT), Robustly optimized BERT approach (RoBERTa), ULMFiT and Extra-Long Network (XLNet), representing a diverse set of deep learning frameworks. The research unfolded through four distinct stages. The initial stage, Data Validation, involved the authors dividing the Amazon dataset, sourced from GitHub, into training and testing subsets, labeled as 0 or 1 to denote negative or positive reviews. Following this, the Statistical Analysis stage ensured the avoidance of overfitting, a critical consideration to prevent misleading results. Moving forward, the third stage involved Exploratory Data Analysis (EDA), where the authors employed TF-IDF and K-means clustering to sift through the data and eliminate irrelevant words, enhancing the precision of the analysis. The culmination of these stages led to the Approach stage, where the authors applied the proposed models to discern their respective performances. Notably, the RoBERTa model emerged as a standout performer, achieving an impressive overall classification accuracy of 82%.

Prakash and Aloysius (2021) introduced a noteworthy contribution to sentiment analysis on tweets by proposing a lexicon-based approach. Their endeavor aimed at tackling the pivotal challenges in sentiment analysis, particularly focusing on optimizing the performance of lexicon-based methodologies.

Tang *et al.* (2016). proposed a deep learning-based sentiment analysis model for product reviews that incorporates attention mechanisms to focus on sentiment-bearing words and phrases, improving the accuracy of sentiment classification. The authors developed a sentiment analysis model based on convolutional neural networks (CNNs) with attention mechanisms. CNNs are effective at extracting local patterns from text, while attention mechanisms allow the model to focus on the most important parts of a sentence for sentiment classification. The model first extracts local features from product reviews using CNNs. Then, an attention mechanism is applied to assign weights to these features based on their importance for sentiment classification. The final sentiment polarity of the review is determined based on the weighted combination of the extracted features.

Vangheese and Nellasivan (2023) proposed a sentiment analysis of tweets concerning food delivery services employing lexicon-based approaches. The primary objective was to discern consumer perceptions towards two prominent food delivery brands, namely Swiggy and Zomato. Utilizing the R programming language, the authors collected data directly from the two brands' Twitter feeds for analysis. The collected data underwent thorough cleaning and pre-processing using various R language techniques to ensure data integrity. Subsequently, the authors employed a lexicon-based method for emotion classification, identifying terms such as "positive," "anger," and "joy" to categorize the sentiments expressed in the tweets. Post-analysis, it was revealed that the Swiggy brand outperformed Zomato, garnering a higher number of positive tweets. This outcome underscores the effectiveness of the lexicon-based approach in gauging and comparing consumer sentiments in the context of food delivery services.

III. MATERIALS AND METHODS

Data Collection

Two separate datasets were collected: the first dataset being a labeled dataset from yelp.com that was obtained from Kaggle.com a reputable website and an excerpt from reviews attained from amazon.com served as the second dataset. The two datasets were merged into one dataset called “training_dataset.csv” for the purpose of this research (Fig.1).

Objectives of the Design

In order to achieve the development of a product review sentiment analyzer, the following objectives were carried out:

a) design a sentiment analyzer using Gaussian Naïve bayes model.
b) to preprocess the dataset using lemmatization and tokenization methods of NLP.
c) to extract and transform features using CountVectorizer.
d) design a user-friendly interface that can capture both individual and batch sentiments using Streamlit.
e) evaluate the performance of the model using accuracy, precision, recall and F1-score.

Fig. 2 shows the model of the product review sentiment analyzer.

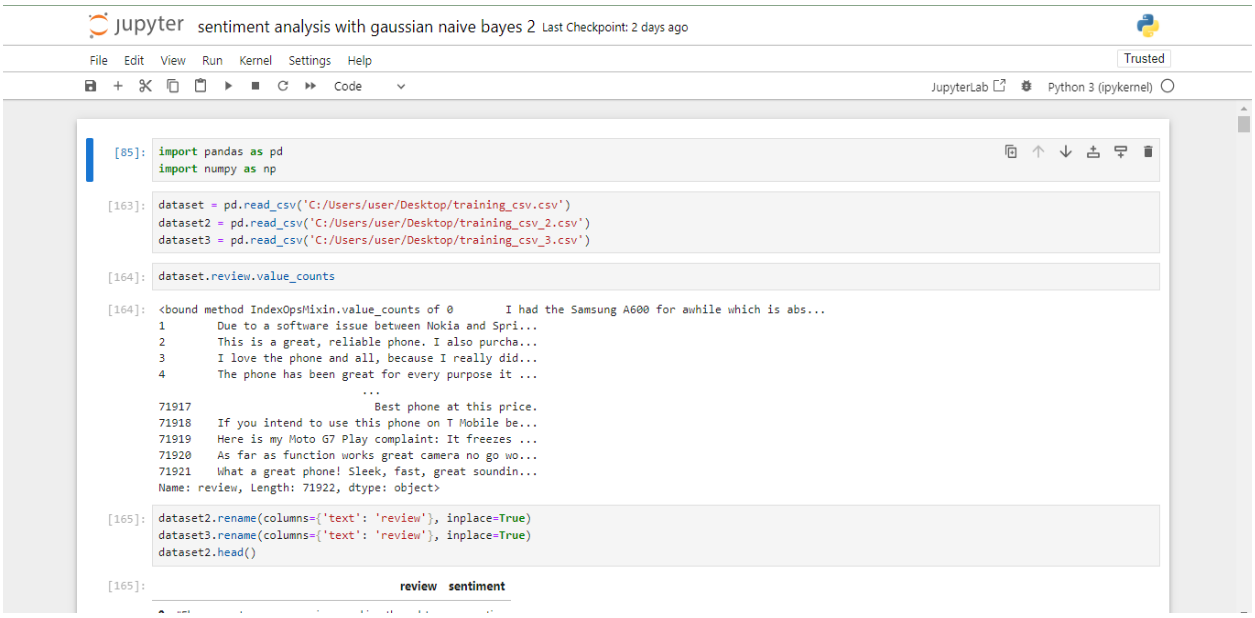

Fig. 1: Merging of the two datasets obtained from kaggle.com

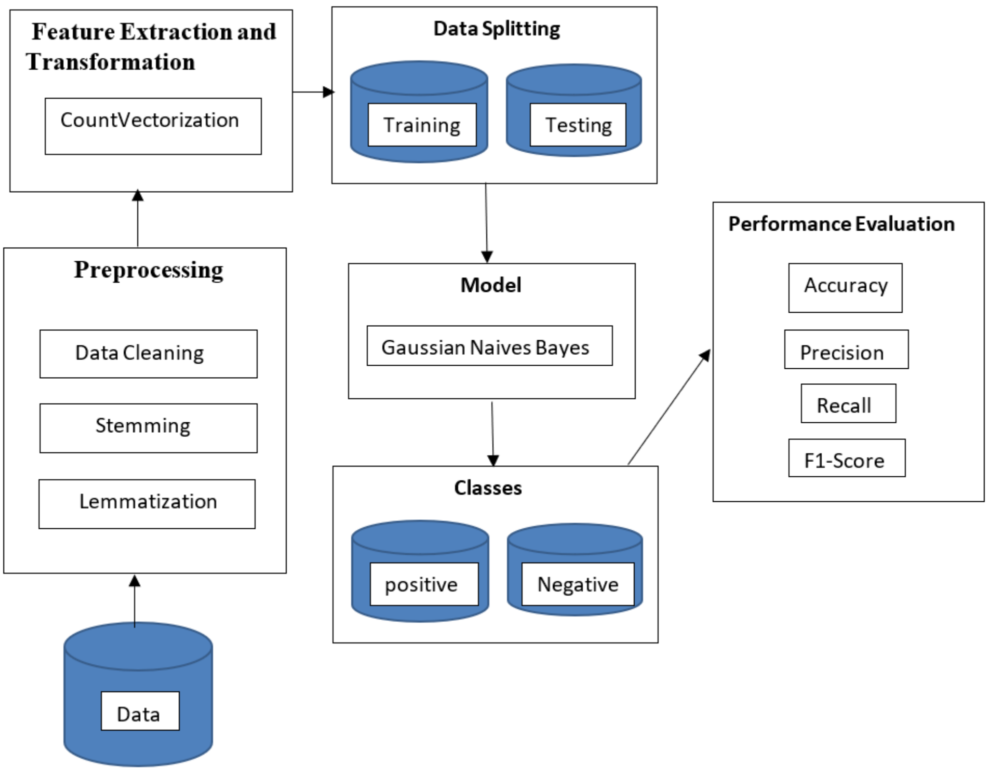

Fig. 2: Product Review Sentiment Analyzer

Preprocessing

Data in its unprocessed form might contain some features which are not relevant to the research. The first stage of preprocessing was cleaning of the dataset whereby the columns of the data were renamed for easy interpretation as seen in Fig. 3 and removing duplicate data, handling null values, and cleaning non-alphabetic characters, as well as converting text to lowercase were implemented to reduce bias and inconsistencies which may affect the performance of the model.

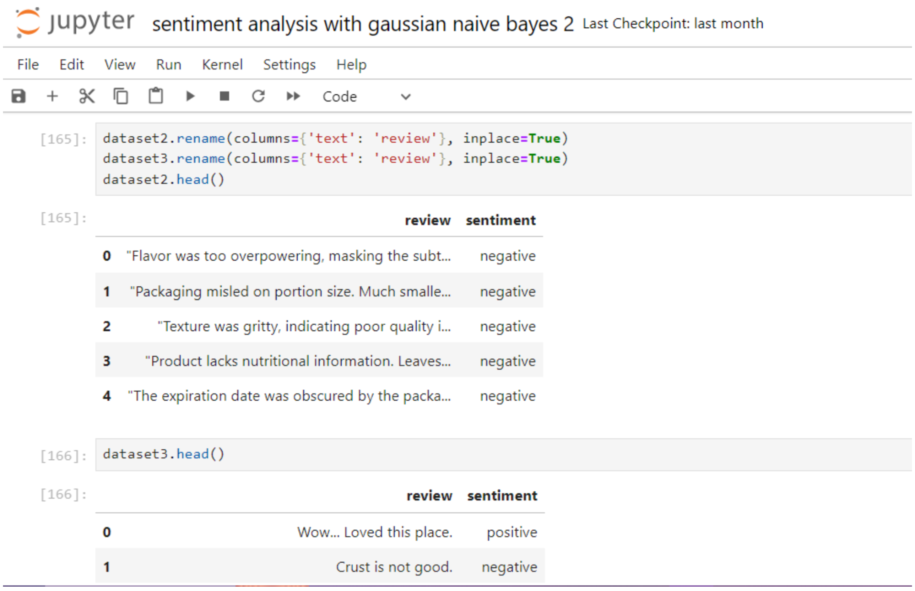

Fig.3: Code block of cleaning of the dataset

The preprocessing techniques such Lemmatization and Tokenization were applied to reduce words to its base or root form and convert a sentence into a list of words respectively as shown in Fig. 4.

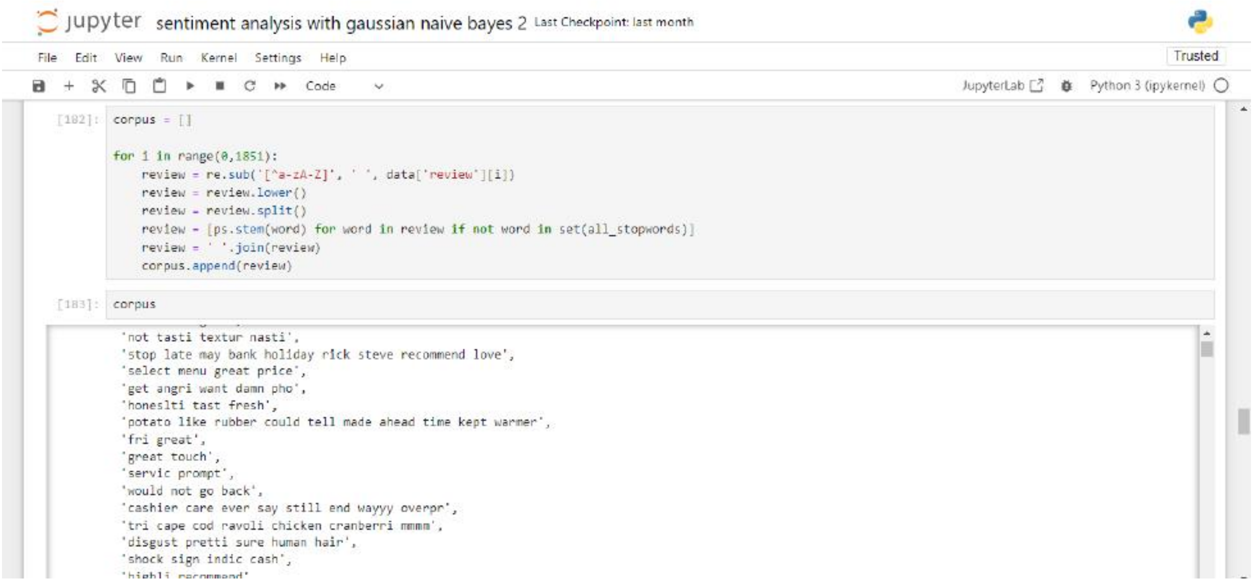

Fig. 4: Code Block for Data Preprocessing

Data Splitting

For effective model evaluation, a randomized split with a fixed random state was used to split the dataset into training and testing sets in the ratio of 80%:20% of data for each respectively (Fig. 5). This approach ensures reproducibility.

Fig. 5: Data Splitting Code Block

Feature Extraction

CountVectorizer was employed to convert raw text into a numerical format suitable for model's training. This technique created a matrix of token counts, where each row represented a document and each column a unique word (Fig. 6). This transformation facilitated machine learning model comprehension of textual data. It allowed extraction of meaningful patterns from text, aligning with the numerical nature of machine learning models and captured word frequencies for sentiment determination. Figs. 7a and b show word cloud for positive and negative reviews.

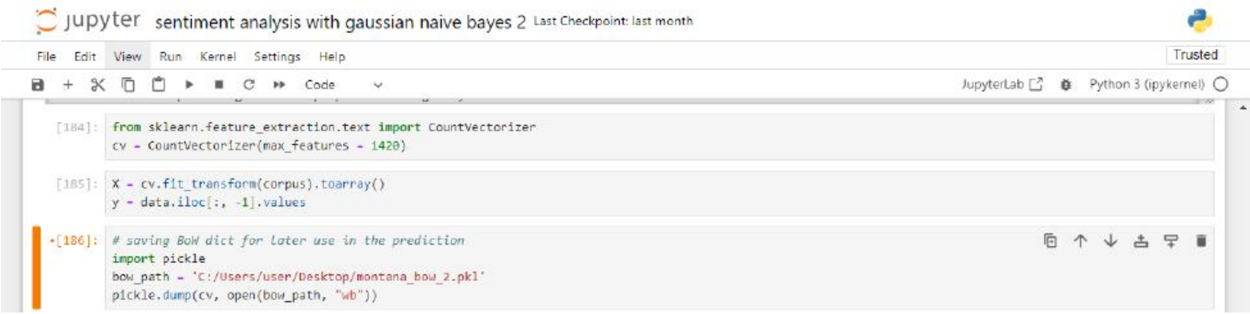

Fig.6: Feature Extraction using CountVectorizer

Fig. 7a: WorldCloud for Positive Reviews

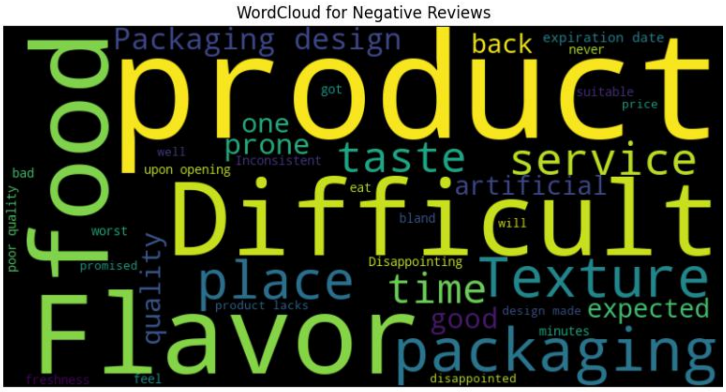

Fig. 7b: WorldCloud for Negative Reviews

Model Selection

The Gaussian Naive Bayes (GNB) was adopted. The choice was motivated by its suitability for handling numerical features which aligns with the transformed data obtained from features extractions using techniques like CountVectorizer and particularly effective in scenarios where features follow a Gaussian distribution.

Training

The model was trained using the training dataset. The process involved exposing the model to the labeled dataset, thereby, allowing it to learn patterns and relationships between the numerical features and sentiment labels (positive and negative). For the testing, the testing dataset was used to test the model.

Performance Evaluation

The model was evaluated using accuracy, precision, recall, and F1-score as the evaluation metrics to gauge the model's effectiveness in classifying sentiments.

Precision: This is used to determine the exactness of the measurement. It measures the accuracy of the positive predictions. It is calculated as:

$$Precision = \frac{True\ Positives}{(True\ Positives + False\ Positives)} \quad (1)$$

Recall: This measures the completeness of positive predictions, that is, measure of how well a model correctly identifies True Positives.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (2)$$

F1-Score: A measure of a model's accuracy on a dataset. It is a harmonic means of both precision and recall of the model. It is determined as:

$$F1_{Score} = 2\ X\ \frac{(precision\ x\ recall)}{(precision + recall)} \quad (3)$$

IV. RESULTS AND DISCUSSION

For the user interface, Fig. 8 shows the user interface featuring the analysis of a customer's review. The user interface accepts both individual and batch sentiments depending on the input at hand.

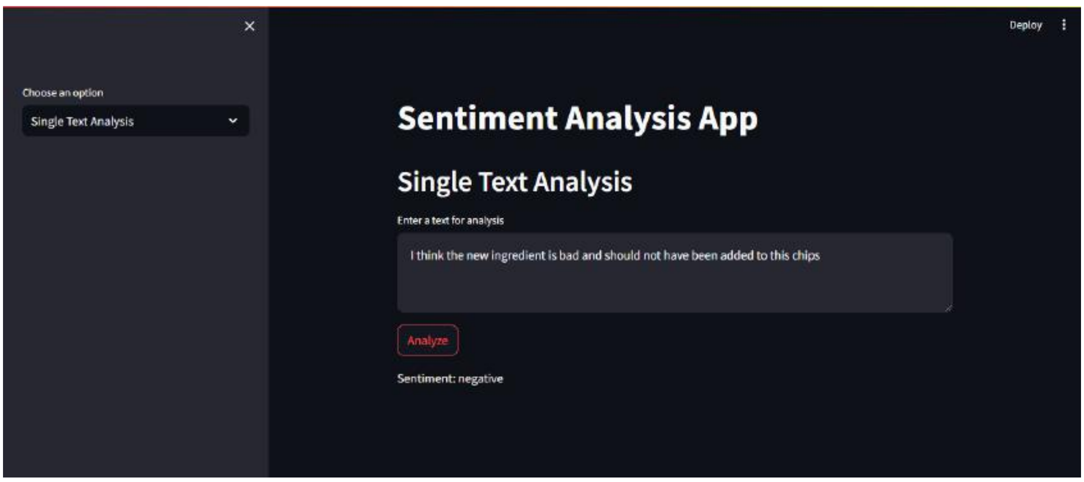

Fig. 8: A User-friendly interface that captures both individual and batch sentiments

For batch analysis, Fig. 9 shows the sample of customers' analyzed sentiments.

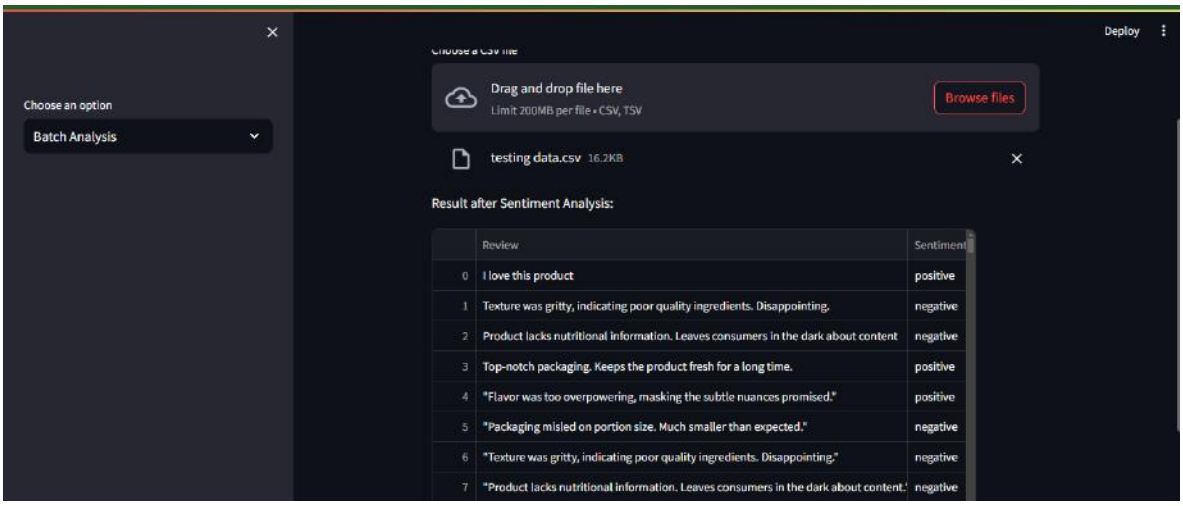

Fig. 9: Sample of analyzed sentiments

The evaluation of the performance of the model returned 90% as accuracy (Fig. 10). This shows the effectiveness and reliability of the model in analyzing sentiments of customers for the products received. The results of the distribution and proportion of the positive and negative reviews are presented on Figs. 11 and 12 respectively. The accuracy of the model returned 90% to show its efficiency.

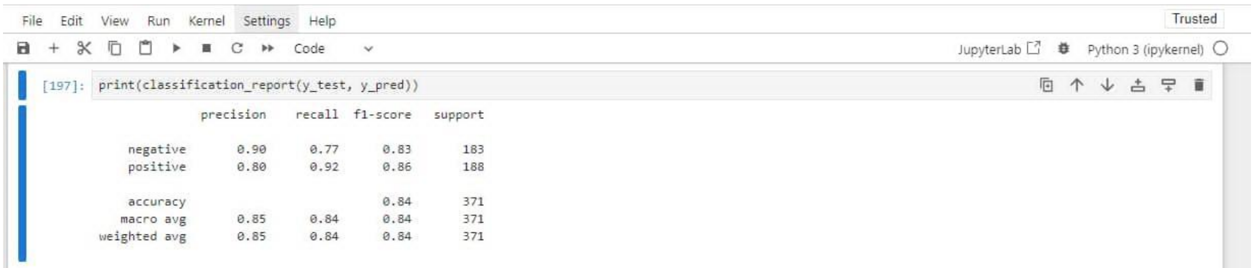

Fig. 10: Performance Evaluation of the model

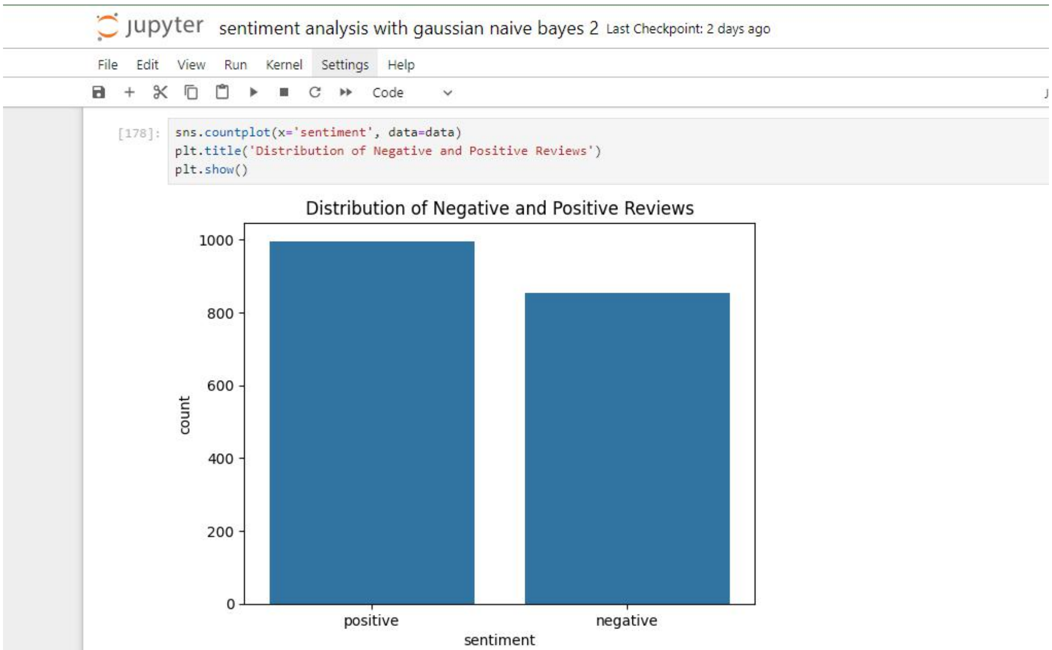

Fig. 11: Bar Chart showing the Distribution of Negative and Positive Reviews

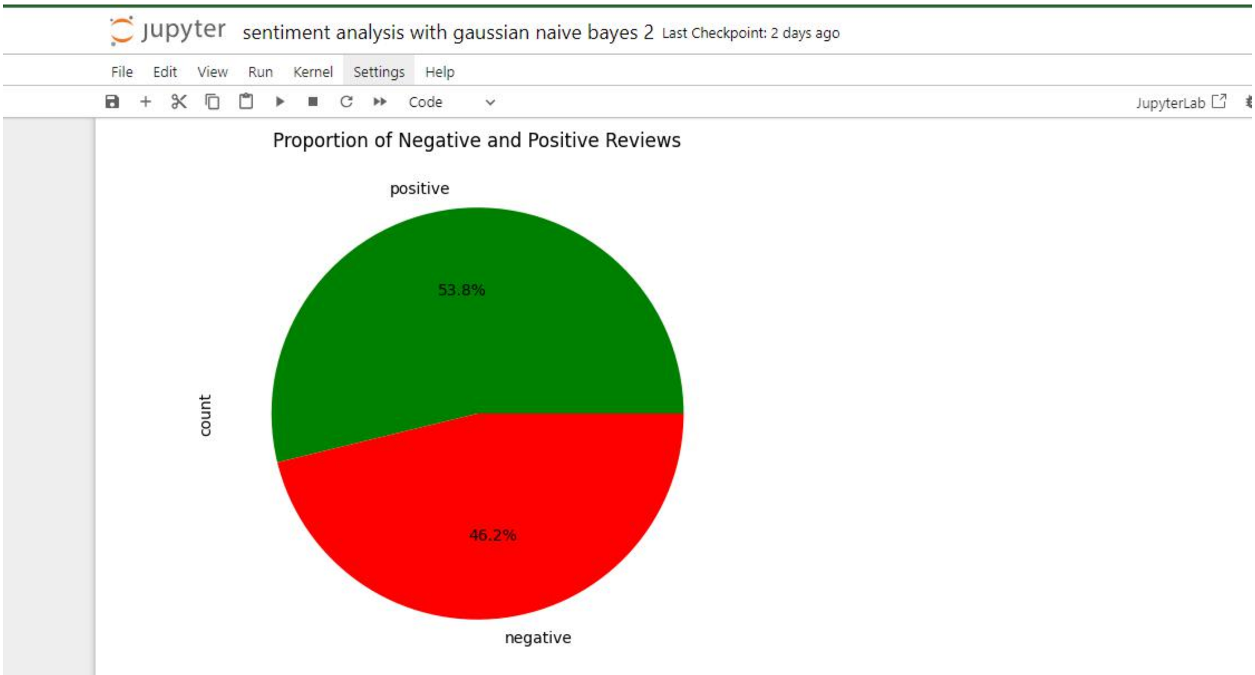

Fig. 12: Pie Chart of the Proportion of Negative and Positive Reviews

V. CONCLUSION

The research shows the feasibility and effectiveness of utilizing Gaussian Naive Bayes in sentiment analysis for e-commerce product review. The user interface enhances accessibility, enabling users to analyze sentiments for both individual and across large datasets. The outcomes of this research contributed to the advancement of sentiment analysis techniques in the context of e-commerce, thereby, providing valuable insights into customer preferences and sentiments.

REFERENCES

[1] Alharbi, .N. M., Alghamdi, .N. S., Alkhammash, .E. H., & Al Amri, .J. F. (2021). Evaluation of Sentiment Analysis via Word Embedding and RNN Variants for Amazon Online Reviews, Mathematical Problems in Engineering, Vol.2021, pp.1-10.

[2] Biyani., A. (2023) A Complete Guide to Sentiment Analysis. Retrieved from https://careerfoundry.com/en/blog/data-analytics/sentiment-analysis/

[3] Korolov, M. (2021) What is sentiment analysis? Using NLP and ML to Extract Meaning. Retrieved from https://www.cio.com/article/189218/what-is-sentiment-analysis-using-nlp-and-ml-to-extract-meaning.html

[4] Lutkevich, B., Chui, W. & Holsk, .B (2023). What is E-commerce? Definition and meaning. Retrieved from https://www/techtarget.com/searchcio/definition/e-commerce

[5] Needle, F. (2023) The Benefits of customer feedback, According to Experts. Retrieved from https//blog.hubspot.com/service/benefits-of-customer-feedback

[6] Noriega, S., Alvarez, H. V., Ramírez, C. E. M., & Cantu-Ortiz, F. J. (2023). Sentiment Analysis of Amazon Reviews using Deep Learning Techniques. Noriega, S., Alvarez, H. V., Ramírez, C. E. M., & Cantu-Ortiz, F. J. (2023). Sentiment Analysis of Amazon Reviews using Deep Learning Techniques.

[7] Oracle. (n.d). What Is Natural Language Processing (NLP)? Retrieved from https://www.oracle.com/ng/artificial-intelligence/what-is-natural-language-processing/

[8] Post., J. (2023). Top E-commerce challenges facing SMBs. Retrieved from https://www.businessnewsdaily.com/6028-small-ecommerce-challenges.html2

[9] Prakash, T. N., & Aloysius, .A. (2021). Textual Sentiment Analysis using Lexicon Based Approaches, Annals of the Romanian Society for Cell Biology, Vol. 25, Issue 4, pp. 9878-9885.

[10] Tang, .D., Wei, .F., Qin, .B., Yang, .N., Liu, .T. & Zhou, .M. (2016). Sentiment Embeddings with Applications to Sentiment Analysis, IEEE Transactions on Knowledge and Data Engineering, Vol. 28, no 2, pp.496-509

[11] Tyagi, N. (2020, April 24). 6 major branches of Artificial Intelligence (AI). Retrieved from https://www.analyticssteps.com/blogs/6-major-branches-artificial-intelligence-ai

[12] Vargheese, M., & Nallasivan, G. (2023). Lexicon-based sentiment analysis of tweets on food delivery services, International Journal of Innovative Research in Engineering and Multidisciplinary Physical Sciences, Vol.11, Issue 2, pp.1-6

[13] Ziyad, S. (2019) Artificial Intelligence Definition, Ethics and standards. The British University in Egypt.